

# The Legal Translating Service

*A Division of Linguistic Systems, Inc.*

P.O. Box 390031 • 130 Bishop Richard Allen Drive • Cambridge, Massachusetts 02139 • Telephone 617-864-3900

## Certification of Translation

COMMONWEALTH OF MASSACHUSETTS  
COUNTY OF MIDDLESEX

On this day of April 24, 2001

Warner V. Heinz

of The Legal Translating Service, a division of Linguistic Systems, Inc., 130 Bishop Richard Allen Drive, Cambridge, Massachusetts 02139, being duly sworn, declared that the attached translation has been made faithfully of his own knowledge by himself and that the attached translation is a true and correct English version of the original German document, to the best of his knowledge and belief.

His qualifications as translator include familiarity with German as a native language and with English, French, and Spanish as acquired languages, and with said languages as languages of instruction and use for more than 35 years, and that he received a Bachelor of Science as well as a Masters in Business Administration from Ludwig Maximilian University in Munich, Germany and that he is employed as a freelance translator by Linguistic Systems, Inc.

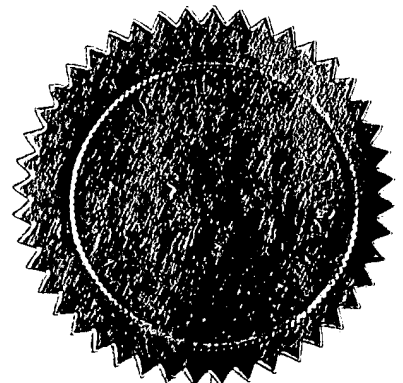
My commission expires March 28, 2008

Translator

*Warner V. Heinz*

*Hugh McAden Oechler*

Hugh McAden Oechler  
Notary Public



Verfahren zur Elimination von Störsignalanteilen in einem Eingangssignal eines auditorischen Systems, Anwendung des Verfahrens und ein Hörgerät

5

Die vorliegende Erfindung betrifft ein Verfahren zur Elimination von Störsignalanteilen in einem Eingangssignal eines auditorischen Systems, eine Anwendung des Verfahrens zum Betrieb eines Hörgerätes sowie ein Hörgerät.

10

Hörgeräte werden bekanntlich bei Personen mit einem geschädigten Gehör eingesetzt, wobei es zunächst das Ziel ist, den Hörschaden möglichst vollständig zu kompensieren. Die Akzeptanz eines Hörgerätes ist beim Hörgeräteträger

15

insbesondere dann erhöht, wenn das Hörgerät auch in einer Umgebung mit starken Störgeräuschen seine Aufgaben zufriedenstellend erfüllt, mit anderen Worten, wenn die Sprachverständlichkeit auch bei starken Störsignalen für den Hörgeräteträger hoch ist.

20

Wenn im folgenden der Begriff "Hörgerät" verwendet wird, so sind darunter sowohl sogenannte Hörhilfen, welche zur Korrektur eines geschädigten Hörvermögens einer Person eingesetzt werden, als auch alle anderen akustischen

25

Kommunikationssysteme, wie zum Beispiel Funkgeräte, zu verstehen.

Zur Verbesserung der Sprachverständlichkeit mit Hilfe von Hörgeräten bei Vorhandensein von Störsignalen sind drei

30

Techniken bekannt:

Als erstes wird auf Hörgeräte verwiesen, welche mit der sogenannten Richtmikrophon-Technologie ausgestattet sind. Diese Technologie erlaubt es, eine räumliche Filterung

35

vorzunehmen, so dass Störgeräusche, welche aus einer anderen Richtung als ein Nutzsignal kommen - beispielsweise

von hinten oder von der Seite - abgeschwächt oder sogar eliminiert werden können. Das auch etwa unter dem Begriff "Beamforming" bekannte Verfahren erfordert mindestens zwei Mikrophone im Hörgerät. Einer der Hauptnachteile bei diesen  
5 bekannten Hörgeräten ist die Tatsache, dass Störgeräusche, welche aus der gleichen Richtung kommen wie das Nutzsignal, nicht abgeschwächt, geschweige denn eliminiert werden können.

10 Des weiteren wird auf die bekannte Möglichkeit verwiesen, mit Hilfe einer Sendeeinheit ein interessierendes Nutzsignal möglichst am Ort des Entstehens aufzunehmen und über eine drahtlose Verbindung direkt in eine  
Empfangseinheit im Hörgerät zu übermitteln. Damit wird  
15 vermieden, dass Störsignale ins Hörgerät gelangen. Dieses in der Hörgeräteindustrie auch etwa unter dem Begriff "FM- (Frequency Modulation)-Technologie" bekannte Verfahren erfordert Zusatzgeräte wie die bei der Schallquelle zu platzierende Sendeeinheit und die mit dem Hörgerät zu  
20 koppelnde Empfangseinheit. Die Handhabung ist für den Hörgerätebenutzer entsprechend umständlich.

Schliesslich wird an dritter Stelle auf Hörgeräte hingewiesen, in denen die Eingangssignale unter Anwendung  
25 von Signalverarbeitungsalgorithmen verarbeiten werden mit dem Ziel, die im Eingangssignal enthaltenen Störsignale zu unterdrücken, zumindest aber abzuschwächen, bzw. die entsprechenden Nutzsignalanteile zu verstärken (sogenanntes "noise canceling"). Hierzu werden die im Eingangssignal  
30 vorhandenen Störsignalanteile in mehreren Frequenzbändern geschätzt, wobei in der Folge allfällige Störsignalanteile vom Eingangssignal des Hörgerätes zur Erzeugung des Nutzsignals subtrahiert werden oder die Störanteile abgeschwächt werden. Dieses Vorgehen wird auch etwa als  
35 "spectral subtraction" bezeichnet. Aus der europäischen Patentschrift mit der Nummer EP-B1-0 534 837 ist ein

solches Verfahren bekannt, das zu annehmbaren Ergebnissen führt. Die Methode der spektralen Subtraktion funktioniert allerdings nur dann gut, wenn die Störsignalanteile bandbegrenzt und stationär sind. Ist diese Annahme nicht erfüllt, wie z.B. bei nicht-stationären Störsignalanteilen, können das Nutzsignal (Sprachsignal, welches nicht-stationär ist) und die Störsignalanteile nicht unterschieden werden. Die spektrale Subtraktion funktioniert in diesem Fall schlecht und die Sprachverständlichkeit ist durch die nicht eliminierten Störgeräuschanteile stark reduziert. Darüber hinaus kann durch die Anwendung der spektralen Subtraktion auch das Nutzsignal beeinträchtigt werden.

Des weiteren hat eine Untersuchung von Bear et. al. ("Spectral Contrast Enhancement of Speech in Noise for Listeners with Sensorineural Hearing Impairment: Effects on Intelligibility, Quality, and Response Times", Journal of Rehabilitation Research and Development 30, Seiten 49 bis 72) gezeigt, dass die Verstärkung des spektralen Kontrastes (spectral enhancement) zwar zu einer subjektiv besseren Qualität des Signals sowie zu einer verminderten Höranstrengung führt aber nicht allgemein zu einer Verbesserung der Sprachverständlichkeit. In diesem Zusammenhang wird auf einen Aufsatz von Frank et. al. mit dem Titel "Evaluation of Spectral Enhancement in Hearing Aids, Combined with Phonemic Compression" (Journal of the Acoustic Society of America 106, Seiten 1452 bis 1464) verwiesen.

Des weiteren wird der Vollständigkeit halber auf die folgenden Dokumente verwiesen:

- T. Baer, B. C. J. Moore "Evaluation of a Scheme to Compensate for Reduced Frequency Selectivity in Hearing -Impaired Subjects", veröffentlicht in

"Modeling Sensorneural Hearing Loss" durch W. Jesteadt, Lawrence Erlbaum Associated, Publishers, Mahwah, New Jersey, 1997;

- 5 • V. Hohmann, "Binaural Noise Reduction and a Localization Model Based on the Statistics of Binaural Signal Parameters", International Hearing Aid Research Conference, Lake Tahoe, 2000;
- US-5 727 072;
- 10 • N. Virag, "Speech enhancement based on masking properties of the human auditory system", Ph.D. Thesis, Ecole Polytechnique Fédérale de Lausanne, 1996.
- WO 91/03042

15 Der vorliegenden Erfindung liegt daher die Aufgabe zugrunde, ein Verfahren anzugeben, bei dem die Elimination von Störsignalanteilen verbessert ist.

20 Diese Aufgabe wird durch die im Patentanspruch 1 angegebenen Massnahmen gelöst. Vorteilhafte Ausgestaltungen der Erfindung sowie eine Anwendung des Verfahrens und ein Hörgerät sind in weiteren Ansprüchen angegeben.

25 Das erfindungsgemässe Verfahren, bestehend aus Signalanalysephase und Verarbeitungsphase, erlaubt es, aus beliebigen Eingangssignalen ein allfälliges Nutzsignal zu extrahieren und unerwünschte Störanteile gezielt auszuschalten bzw. erwünschte Signalanteile neu zu generieren. Damit kann eine viel bessere, der  
30 Umgebungssituation angepasste Störgeräuschunterdrückung erreicht werden. Im Gegensatz zum herkömmlichen Noise Canceling werden mit dem erfindungsgemässen Verfahren die Nutzsignale nicht beeinträchtigt. Des weitern können auch nicht-stationäre Störgeräusche aus dem Eingangssignal  
35 entfernt werden. Schliesslich wird darauf hingewiesen, dass

mit herkömmlichen Störgeräusch-Unterdrückungsalgorithmen es nicht möglich ist, eine Synthese des Nutzsignals vorzunehmen.

- 5 Die Erfindung wird nachfolgend anhand von Zeichnungen beispielsweise näher erläutert. Dabei zeigt

Fig. 1 eine schematische Darstellung des  
erfindungsgemässen Verfahrens anhand eines  
10 Blockdiagramms,

Fig. 2, wiederum in schematischer Darstellung, einen Teil  
des Blockdiagramms gemäss Fig. 1 und

15 Fig. 3 eine weitere Ausführungsform des in Fig. 2  
dargestellten Teilblockdiagramms.

In Fig. 1 ist ein Blockdiagramm dargestellt, anhand dessen  
das erfindungsgemässe Verfahren, bestehend aus einer  
20 Signalanalysephase I und einer Signalverarbeitungsphase II,  
beschrieben wird. In der Signalanalysephase I wird ein  
einem auditorischen System beaufschlagtes Eingangssignal  
ES, welches sowohl Stör- als auch Nutzsignalanteile SS bzw.  
NS aufweisen kann, anhand von auditorisch-basierten  
25 Prinzipien, was im folgenden noch erläutert wird,  
analysiert. In der Folge wird in der  
Signalverarbeitungsphase II eine Störgeräuschbefreiung  
vorgenommen, wobei die in der Signalanalysephase I  
gewonnenen Erkenntnisse über die Störsignalanteile SS und  
30 über die Nutzsignalanteile NS dazu verwendet werden.  
Grundsätzlich werden hierzu zwei Ausführungsvarianten  
vorgeschlagen: Die erste besteht darin, dass zur Gewinnung  
des bzw. der Nutzsignale NS die unerwünschten  
Störsignalanteile SS aus dem Eingangssignal ES entfernt  
35 werden, bzw. dass die unerwünschten Störsignalanteile SS  
unterdrückt oder abgeschwächt werden. Bei der zweiten

Methode wird mittels einer Synthese das Nutzsignal NS bzw. NS' erzeugt.

5 In einer weiteren Ausführungsform des erfindungsgemässen  
Verfahrens ist darüber hinaus auch vorgesehen, beide  
vorstehend genannten Vorgehensweisen zu berücksichtigen,  
mit anderen Worten, es wird sowohl die Unterdrückung der  
ermittelten Störsignalanteile SS als auch die Synthese des  
identifizierten Nutzsignals NS bzw. NS' in Kombination  
10 angewendet.

Im Unterschied zu herkömmlichen Verfahren zur  
Geräuschunterdrückung, bei denen ein Eingangssignal -  
ebenfalls in einer Signalanalysephase - lediglich aufgrund  
15 der Stationarität bzw. Nicht-Stationarität untersucht wird,  
basiert das erfindungsgemässe Verfahren auf einer  
auditorisch-basierten Signalanalyse. Dabei werden zumindest  
auditorisch-basierte Merkmale, wie Lautheit, spektrales  
Profil (timbre), harmonische Struktur (pitch), gemeinsame  
20 Ein- und Ausschwingzeiten (on-/offset), kohärente  
Amplituden- und Frequenzmodulation, kohärente Phasen,  
interaurale Laufzeit- und Pegelunterschiede und andere, aus  
dem Eingangssignal ES extrahiert, wobei die Extraktion  
einzelner oder aller Merkmale vorgesehen sein kann. Die  
25 Definition und weitere Angaben zu auditorischen Merkmalen  
sind der druckschriftlichen Veröffentlichung von A. S.  
Bregman mit dem Titel "Auditory Scene Analysis" (MIT Press,  
Cambridge, London, 1990) entnehmbar. Es wird darauf  
hingewiesen, dass das erfindungsgemässe Verfahren nicht auf  
30 die Extraktion von auditorisch-basierten Merkmalen  
eingeschränkt ist. Vielmehr ist denkbar, und soll im  
weiteren als zusätzliche, vorteilhafte Ausgestaltung des  
erfindungsgemässen Verfahrens verstanden werden, dass neben  
auditorisch-basierten Merkmalen zusätzlich auch rein  
35 technisch-basierte Merkmale - wie zum Beispiel  
Nulldurchgangsraten, zeitliche Pegelschwankungen,

verschiedene Modulationsfrequenzen, spektraler Schwerpunkt, Amplitudenverteilung, u.a. - extrahiert werden.

5 In einer spezifischen Ausführungsform ist vorgesehen, dass die Extraktion der Merkmale entweder aus dem Zeitsignal oder in verschiedenen, Frequenzbändern vorgenommen wird. Dazu kann eine gehörgerechte Filterbank (E. Zwicker, H. Fastl, Psychoacoustics -- Facts and Models , Springer Verlag, 1999) oder auch eine technisch basierte Filterbank  
10 wie z. B. eine FFT- oder Wavelet-Filterbank verwendet werden.

Die Auswertung der ermittelten Merkmale, seien dies nun auditorisch-basierte oder technisch-basierte Merkmale,  
15 erlaubt die Identifikation und Unterscheidung verschiedener Signalanteile  $SA_1$  bis  $SA_n$ , wobei einzelne dieser Signalanteile  $SA_1$  bis  $SA_n$  interessierende Nutzsignale NS oder zu eliminierende Störsignale SS sind.

20 Die Auftrennung in die Signalanteile  $SA_1$  bis  $SA_n$  wird erfindungsgemäss mit Hilfe von zwei verschiedenen Ansätzen erreicht. Die beiden Ansätze werden anhand der Figuren 2 und 3 im folgenden erläutert.

25 In Fig. 2 sind die in der Signalanalysephase I ablaufenden Verfahrensschritte in einem Blockdiagramm dargestellt. Es handelt sich dabei um zwei in Serie geschaltete Einheiten, nämlich um eine Merkmalsextraktionseinheit 20 und um eine Gruppierungseinheit 21.

30 In der Merkmalsextraktionseinheit 20 erfolgt die bereits erläuterte Extraktion von auditorisch-basierten und gegebenenfalls auch von technisch-basierten Merkmalen  $M_1$  bis  $M_j$  zur Charakterisierung des Eingangssignals ES. Diese  
35 Merkmale  $M_1$  bis  $M_j$  werden in der Folge in der Gruppierungseinheit 21 mit Hilfe der Methode der primitiven

Gruppierung, welche von A. S. Bregman mit dem Titel  
"Auditory Scene Analysis" (MIT Press, Cambridge, London,  
1990) beschrieben worden ist, geordnet. Diese an sich  
bekannte Methode ist kontext-unabhängig und basiert auf der  
5 sequentiellen Abfolge verschiedener Arbeitsschritte, mit  
welchen das Eingangssignal ES auf Grund der extrahierten  
Merkmale  $M_1$  bis  $M_j$  in verschiedene, den unterschiedlichen  
Klangquellen zugeordneten Signalanteilen  $SA_1$  bis  $SA_n$   
unterteilt wird. Dieser Ansatz wird auch etwa als "bottom-  
10 up" oder als "data-driven" bezeichnet. Diesbezüglich wird  
auf die Veröffentlichung von G. Brown mit dem Titel  
"Computational Auditory Scene Analysis: A Representational  
Approach" (Ph.D. Thesis, University of Sheffield, 1992) und  
auf die Veröffentlichung von M. Cooke mit dem Titel  
15 "Modelling Auditory Processing Analysis and Organisation"  
(Ph.D. Thesis, University of Sheffield, 1993) verwiesen.  
Eine bevorzugte Ausführungsvariante ist in Fig. 3, wiederum  
in einem Blockschaltbild, dargestellt, wobei die Verwendung  
der Methode der Schema-basierten Gruppierung vorgesehen  
20 ist, welche wiederum von A. S. Bregman (a.a.O.) ausführlich  
erläutert worden ist. Die Methode der Schema-basierten  
Gruppierung ist kontext-abhängig und wird auch etwa als  
"top-down" oder "prediction-driven" bezeichnet.  
Diesbezüglich sei auf die Veröffentlichung von D. P. W.  
25 Ellis mit dem Titel "Prediction-Driven Computational  
Auditory Scene Analysis" (Ph.D. Thesis, Massachusetts  
Institute of Technology, 1996) verwiesen.

Wie aus Fig. 3 ersichtlich ist, wird neben der  
30 Merkmalsextraktionseinheit 20 und der Gruppierungseinheit  
21 zusätzlich eine Hypotheseeinheit 22 in der  
Signalanalysephase I tätig. Dabei wird bereits aufgrund der  
in Fig. 3 dargestellten Struktur deutlich, dass nicht mehr  
nur eine sequentielle Abfolge von Arbeitsschritten  
35 vorgesehen ist, sondern es wird, unter Berücksichtigung von  
einem Vorwissen V, welche der Hypotheseeinheit 22 zugeführt

wird, eine Hypothese H über die Art des Eingangssignals ES aufgrund der extrahierten Merkmale  $M_1$  bis  $M_j$  und aufgrund der Signalanteile  $SA_1$  bis  $SA_n$  erzeugt. Aufgrund der Hypothese H wird vorzugsweise sowohl die Merkmalsextraktion  
5 in der Merkmalsextraktionseinheit 20 als auch die Gruppierung der Merkmale  $M_1$  bis  $M_j$  einer momentanen Situation angepasst. Die Hypothese H wird also mittels einer "bottom-up"-Analyse sowie aufgrund des Vorwissens V über den akustischen Kontext generiert. Die Hypothese H  
10 legt wiederum den Kontext der Gruppierung fest und basiert auf Kenntnissen und Annahmen über die akustische Umgebung sowie auf der Gruppierung selbst. Die in der Signalanalysephase I ablaufenden Verfahrensschritte sind also nicht mehr rein sequentiell, sondern es ist eine  
15 Rückkopplungsschleife vorgesehen, welche die Anpassung an die jeweilige Situation ermöglicht.

Mit der eben beschriebenen bevorzugten Ausführungsvariante des erfindungsgemässen Verfahrens ist es zum Beispiel  
20 möglich, die Sprachverständlichkeit bei einem bekannten Sprecher, von dem das Vorwissen beispielsweise aus der Art der Sprache, typischen Pitchfrequenzen, Sprechgeschwindigkeit und Formantenfrequenzen besteht, wesentlich zu verbessern gegenüber der Situation, wo keine  
25 Angaben über den Sprecher mitberücksichtigt werden.

Gemäss dem erfindungsgemässen Verfahren und unter Berücksichtigung der vorstehenden Ausführungen im Zusammenhang mit der Gruppierung können für die Bildung der  
30 auditorischen Objekte, d.h. Signalanteile  $SA_1$  bis  $SA_n$ , in beiden erwähnten Ansätzen für die Gruppierung die Prinzipien der Gestalttheorie (E. B. Goldstein, Wahrnehmungspsychologie, Spektrum Akademischer Verlag, 1996) auf die Merkmale  $M_1$  bis  $M_j$  angewendet werden. Diese  
35 beinhalten insbesondere

- Kontinuität,
- Nähe,
- Ähnlichkeit,
- gemeinsames Schicksal,
- 5 - Geschlossenheit und
- gute Fortsetzung.

Beispielsweise weisen Merkmale, die nicht kontinuierlich und nicht abrupt ändern, auf die Zugehörigkeit zu einer  
10 bestimmten Signalquelle hin. Zeitlich aufeinanderfolgende Merkmale mit ähnlicher harmonischer Struktur (pitch) weisen auf spektrale Nähe hin und werden derselben Signalquelle zugeordnet. Auch andere ähnliche Merkmale, wie zum Beispiel Modulation, Pegel oder spektrales Profil, ermöglichen die  
15 Gruppierung zu einzelnen Klangkomponenten. Ein gemeinsames Schicksal, wie zum Beispiel gemeinsames Ein- und/oder Ausschwingen und kohärente Modulation, deutet ebenfalls auf die Zugehörigkeit zum selben Signalanteil hin. Die Annahme der Geschlossenheit im zeitlichen Ablauf erleichtert die  
20 Interpretation abrupter Änderungen, so trennen Signallücken verschiedene Ereignisse oder Quellen, währenddem überlagerte Anteile auf mehrere Quellen hindeuten.

Unter Fortführung der vorstehenden Ausführungen kann über  
25 dies festgehalten werden, dass sich auch das Kriterium "gute Fortsetzung" für Schlussfolgerungen vorzüglich eignet. So ändert ein Signal normalerweise seinen Charakter nicht plötzlich, weshalb langsame Änderungen demselben Signalanteil zugeordnet werden, währenddem schnelle Wechsel  
30 neue Signalanteile erzeugen.

Weitere Gruppierungsmöglichkeiten ergeben sich aus den extrahierten Merkmalen  $M_1$  bis  $M_5$  selber. So kann durch die Analyse der Lautheit bestimmt werden, ob ein Signalanteil  
35 überhaupt vorhanden ist oder nicht. Des weitern ist das spektrale Profil verschiedener Klangkomponenten

(Signalanteile) in der Regel unterschiedlich und ermöglicht daher die Differenzierung verschiedener auditorischer Objekte. Eine ermittelte harmonische Struktur (pitch) erlaubt ihrerseits Rückschlüsse auf einen tonalen  
5 Signalanteil, der durch Pitch-Filterung erhalten werden kann. Die Übertragungsfunktion eines Pitch-Filters kann dabei beispielsweise die folgende Übertragungsfunktion aufweisen:

10 
$$H_{\text{pitch}}(z) = 1 - z^{-k},$$

wobei  $z^{-k}$  der Periodenlänge der Pitchfrequenz entspricht. Durch Pitch-Filterung können sodann die tonalen  
15 Signalanteile von den restlichen Signalanteilen getrennt werden.

Die Analyse kohärenter Modulationen erlaubt es, spektrale Komponenten zu gruppieren, die mit demselben zeitlichen Muster moduliert sind, bzw. zu trennen, wenn die Muster  
20 unterschiedlich sind. Damit lassen sich insbesondere verschiedene Sprachanteile im Signal identifizieren und in der Folge auch trennen.

Die Auswertung gemeinsamer Ein- und Ausschwingvorgänge  
25 erlaubt es zu erkennen, welche Signalanteile mit verschiedenem Frequenzgehalt zusammengehören. Grössere asynchrone Amplitudenanstiege bzw. -abfälle indizieren wiederum unterschiedliche Signalanteile.

30 Nach der Identifikation der einzelnen Signalanteile  $SA_1$  bis  $SA_n$  in der Signalanalysephase I kann die eigentliche Störgeräuschbefreiung in der Signalverarbeitungsphase II vorgenommen werden (Fig. 1). Eine Ausführungsform des  
erfindungsgemässen Verfahrens besteht darin, die  
35 Störanteile in den Frequenzbändern, in welchen sie auftreten, abzusenken bzw. zu unterdrücken. Das gleiche

Resultat hat die Massnahme, dass identifizierte  
Nutzsignalanteile verstärkt werden. In den gleichen  
erfindungsgemässen Lösungsbereich gehört die Kombination  
von Absenken bzw. Unterdrücken von Störsignalanteilen und  
5 die Verstärkung von Nutzsignalanteilen.

Eine weitere Ausführungsform der in der  
Signalverarbeitungsphase II vorgenommenen  
Verfahrensschritte besteht darin, die gruppierten und als  
10 Nutzsignal identifizierten Signalanteile neu  
zusammenzumischen.

Eine noch weitere Ausführungsform des erfindungsgemässen  
Verfahrens besteht darin, dass das Nutzsignal NS bzw. das  
15 geschätzte Nutzsignal NS mittels der in der  
Signalanalysephase I gewonnenen Informationen neu  
synthetisiert wird. Dabei besteht eine bevorzugte  
Ausführungsvariante darin, dass mittels Analyse der  
harmonischen Struktur (Pitch-Analyse) die verschiedenen  
20 Grundfrequenzen der Nutzsignale extrahiert werden und dass  
die spektralen Pegel der harmonischen Oberwellen  
beispielsweise mittels Lautheits- oder LPC-Analyse bestimmt  
werden ( S. Launer, "Loudness Perception in Listeners with  
Sensorineural Hearing Loss", Dissertation, Universität  
25 Oldenburg, 1995, J. R. Deller, J. G. Proakis, J. H. L.  
Hansen, "Discrete-Time Processing of Speech Signals",  
Macmillan Publishing Company, 1993). Aus dieser Information  
kann dann ein vollständig synthetisch generiertes Signal  
für tonale Sprachanteile erzeugt werden. In Weiterführung  
30 der vorstehend genannten, bevorzugten Ausführungsvariante  
wird vorgeschlagen, eine Kombination von  
Nutzsignalverstärkung und Nutzsignalsynthese vorzusehen.

Das erfindungsgemässe Verfahren, bestehend aus  
35 Signalanalysephase I und Signalverarbeitungsphase II,  
erlaubt es also, aus beliebigen Eingangssignalen ES ein

allfälliges Nutzsignal NS zu extrahieren und unerwünschte Störanteile SS gezielt auszuschalten bzw. erwünschte Signalanteile NS neu zu generieren. Damit kann eine viel bessere, der Umgebungssituation angepasste Geräuschunterdrückung erreicht werden. Im Gegensatz zum herkömmlichen Noise Canceling werden mit dem erfindungsgemässen Verfahren die Nutzsignale nicht beeinträchtigt. Des weitern können auch nicht-stationäre Störgeräusche aus dem Eingangssignal ES entfernt werden. Schliesslich wird darauf hingewiesen, dass mit herkömmlichen Störgeräusch-Unterdrückungsalgorithmen es nicht möglich ist, eine Synthese des Nutzsignals vorzunehmen.

Eine weitere Ausführungsform des erfindungsgemässen Verfahrens besteht darin, dass dieses mit den eingangs erwähnten Methoden wie Beamforming, binaurale Ansätze zur Störgeräuschlokalisation und -Unterdrückung oder durch Klassifikation der akustischen Umgebung und entsprechender Programmwahl kombiniert wird.

Zwei Beispiele ähnlicher Ansätze zur Störgeräuschbefreiung, welche sich aber nur der primitiven Gruppierung bedienen, sind im Folgenden genannt. Unoki und M. Akagi, "A method of signal extraction from noisy signal based on auditory scene analysis", Speech Communication, 27, Seiten 261 bis 279, 1999, und WO 00/01200. In beiden Ansätzen wird die Störgeräuschunterdrückung mittels Extraktion einiger weniger auditorischer Merkmale und kontextunabhängiger Gruppierung durchgeführt. Die hier vorgeschlagene Lösung ist jedoch vollständiger und dem auditorischen System näher. Es wird darauf hingewiesen, dass das erfindungsgemässe Verfahren nicht auf Sprache als Nutzsignal beschränkt ist. Darüber hinaus werden alle bekannten auditorischen Mechanismen sowie technisch basierte Merkmale verwendet. Ausserdem erfolgt die

- 15 -

**Patentansprüche:**

1. Verfahren zur Elimination von Störsignalanteilen (SS) in  
5 einem Eingangssignal (ES), wobei das Verfahren darin  
besteht,

- 10 . dass in einer Signalanalysephase (I) die  
Störsignalanteile (SS) und ein im Eingangssignal (ES)  
enthaltenees Nutzsignal (NS) charakterisiert werden und
- 15 . dass in einer Signalverarbeitungsphase (II) aufgrund  
der in der Signalanalysephase (I) erhaltenen  
Charakterisierung das Nutzsignal (NS) bzw. ein  
geschätztes Nutzsignal (NS') ermittelt bzw. erzeugt  
wird,

wobei die Charakterisierung der Signalanteile (SS, NS)  
zumindest unter Verwendung von auditorisch-basierten  
20 Merkmalen ( $M_1$  bis  $M_j$ ) erfolgt.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass  
ein oder mehrere der nachfolgenden auditorisch-basierten  
Merkmale ( $M_1$  bis  $M_j$ ) zur Charakterisierung der  
25 Signalanteile (NS, SS) verwendet werden: Lautheit,  
spektrales Profil, harmonische Struktur, gemeinsame Ein-  
und Ausschwingzeiten, kohärente Amplituden- und  
Frequenzmodulation, kohärente Phasen, interaurale Laufzeit-  
und Pegelunterschiede.

30

3. Verfahren nach Anspruch 1 oder 2, dadurch  
gekennzeichnet, dass die auditorisch-basierten Merkmale ( $M_1$   
bis  $M_j$ ) in unterschiedlichen Frequenzbändern ermittelt  
werden.

35

145

- 16 -

4. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass die Charakterisierung der Signalanteile (SS, NS) durch Auswertung der in der Signalanalysephase (I) ermittelten Merkmale ( $M_1$  bis  $M_j$ )
- 5 unter Verwendung der Methode der primitiven Gruppierung erfolgt.
5. Verfahren nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, dass die Charakterisierung der
- 10 Signalanteile (SS, NS) durch Auswertung der in der Signalanalysephase (I) ermittelten Merkmale unter Verwendung der Methode der Schema-basierten Gruppierung erfolgt.
- 15 6. Verfahren nach Anspruch 5, dadurch gekennzeichnet, dass über die Art eines Signalanteils (SS, NS) eine Hypothese erstellt bzw. vorgegeben wird, die bei der Gruppierung der ermittelten Merkmale ( $M_1$  bis  $M_j$ ) berücksichtigt wird.
- 20 7. Verfahren nach Anspruch 5 oder 6, dadurch gekennzeichnet, dass zur Charakterisierung der Signalanteile (NS, SS) die auditorisch-basierten und die gegebenenfalls anderen Merkmale ( $M_1$  bis  $M_j$ ) mit den Prinzipien der Gestalttheorie gruppiert werden.
- 25 8. Verfahren nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, dass die als Störsignale (SS) identifizierten Signalanteile unterdrückt und/oder dass die als Nutzsignale (NS) bzw. als geschätzte Nutzsignale (NS')
- 30 identifizierten Signalanteile verstärkt werden.
9. Verfahren nach einem der Ansprüche 1 bis 8, dadurch gekennzeichnet, dass das Nutzsignal (NS) bzw. ein geschätztes Nutzsignal (NS') in der Signalverarbeitungsphase
- 35 (II) aufgrund von in der Signalanalysephase (I) ermittelten Merkmalen ( $M_1$  bis  $M_j$ ) synthetisiert wird.

210

- 17 -

10. Verfahren nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, dass in der Signalanalysephase (I) mit Hilfe einer Analyse der harmonischen Struktur verschiedene Grundfrequenzen der Signalanteile des Nutzsignals (NS) bzw. des geschätzten Nutzsignals (NS') extrahiert werden, dass, insbesondere mit Hilfe einer Lautheits- oder LPC-Analyse, spektrale Pegel von harmonischen Oberwellen dieser Signalanteile bestimmt werden und dass aufgrund der spektralen Pegel und der harmonischen Oberwellen ein Nutzsignal für tonale Sprachanteile synthetisiert wird.

11. Verfahren nach einem der Ansprüche 1 bis 7, dadurch gekennzeichnet, dass in der Signalanalysephase (I) mit Hilfe einer Analyse der harmonischen Struktur nichttonale Signalanteile des Nutzsignals (NS) bzw. des geschätzten Nutzsignals (NS') extrahiert werden, dass, insbesondere mit Hilfe einer Lautheits- oder LPC-Analyse, spektrale Pegel dieser Signalanteile bestimmt werden und dass mit Hilfe eines Noisegenerators ein Nutzsignal für nichttonale Sprachanteile synthetisiert wird.

12. Verfahren nach Anspruch 10 oder 11, dadurch gekennzeichnet, dass das Nutzsignal (NS) bzw. das geschätzte Nutzsignal (NS') verstärkt wird.

13. Anwendung des Verfahrens nach einem der Ansprüche 1 bis 12 zum Betrieb eines Hörgerätes.

14. Hörgerät, das nach dem Verfahren gemäss einem der Ansprüche 1 bis 12 arbeitet.

190

- 18 -

**Zusammenfassung:**

- Verfahren zur Elimination von Störsignalanteilen (SS) in einem Eingangssignal (ES), wobei das Verfahren darin
- 5 besteht, dass in einer Signalanalysephase (I) die Störsignalanteile (SS) und ein im Eingangssignal (ES) enthaltenes Nutzsignal (NS) charakterisiert werden und dass in einer Signalverarbeitungsphase (II) aufgrund der in der
- 10 Signalanalysephase (I) erhaltenen Charakterisierung das Nutzsignal (NS) bzw. ein geschätztes Nutzsignal (NS') ermittelt bzw. erzeugt wird, wobei die Charakterisierung der Signalanteile (SS, NS) zumindest unter Verwendung von auditorisch-basierten Merkmalen ( $M_1$  bis  $M_j$ ) erfolgt.
- 15 Des weiteren ist eine Anwendung des erfindungsgemässen Verfahrens und ein Hörgerät, das nach dem erfindungsgemässen Verfahren arbeitet, angegeben.

(Fig. 1)

100